P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : *https://tci-thaijo.org/index.php/SciTechAsia* **Science & Technology Asia**
Vol. 23 No.4 October - December 2018 Page: [7-17]
Original research article

A Comparative Experiment in Classifying Jewelry Images Using Convolutional Neural Networks

Vishakha Singh* and Phisan Kaewprapha

Department of Electrical and Computer Engineering, Faculty of Engineering, Thammasat University, Pathum Thani 12120, Thailand

Received 4 June 2018; Received in revised form 27 September 2018
Accepted 24 October 2018; Available online 19 December 2018

ABSTRACT

A machine learning approach has been used in this work to categorize jewelry images into five different classes. This classification was achieved by using the convolutional neural network (CNN). The objective was to find different approaches that can be competent for the image classification and recognition. The images used in this work are drawn directly from the jewelry industries and companies. The first technique uses support vector machine along with the features that were extracted from the input images using AlexNet. The second method involves the use of Inception-v3 model for performing the same. Upon experimenting, it was derived that both the approaches performed well, however, Inception-v3 was found to be more successful by 0.9%. The Inception-v3 was then further taken to train the dataset from scratch which resulted in better consistency.

Keywords: Artificial Intelligence (AI); Convolutional Neural Networks (CNN); Image Classification; Support Vector Machine (SVM); Inception-v3

1. Introduction

The world today has become technology dependent and artificial intelligence is playing a crucial role in it. Artificial intelligence (AI) helps to create computer programs that has the ability to learn, modify and evolve by themselves. One of the branches of AI is machine learning that allows the computers to gain information and knowledge without being specifically programmed. They try to impersonate the way our brain processes information [1, 2, 3]. The field of machine learning is enormous, and it is expanding rapidly. It has various specialties, based on which they are continuously divided.

One of those sub-field of machine learning is deep learning [4, 5] which is concerned with the algorithms that are inspired by the human brain structure.

*Corresponding author: vish12150@gmail.com doi: 10.14456/scitechasia.2018.27

Furthermore, the deep learning consists of a feed forward neural network known as convolutional neural network (CNN). CNN works exactly like the regular neural network, the difference is of the "convolution layer", which is provided at the beginning of a CNN. This network, instead of taking an entire image, they break the image into several blocks, which then the machine anticipates what each block is. Ultimately, the system tries to predict what is in the picture based on the prediction of all the blocks. The system detects the input by parallelizing the operations. Convolutional neural networks are very significant in improving the accuracy of image classification.

Image classification can be described as a process where pixels in the image are allocated to various categories of interest. Every electronic image is made up of pixels that has some value or number associated with it. Image classification organizes the data by analyzing that numerical property (maps numbers to symbols).

1.1 Motivation

Jewelry making depends on intensive human labor and because of the tremendous growth in the online shopping sector, almost majority of the people these days are turning towards these online sites for their daily shopping. From groceries to clothes, everything is available and is delivered to our doorstep at the click of the mouse button. Jewelry industry being the top runner of them all, with the global sales of €148 billion is set to touch €250 billion by 2020 [6]. The descriptions available about each jewelry item on different online websites are done manually by the trained workers. Sites such as, amazon.com has 152 categories of "types", 51 categories of "colors", 18 categories of "shapes" and 203 categories of "materials" [7]. The process of categorizing each item manually takes huge amount of time and energy which makes it vulnerable to human errors. Sometimes, people are also sensitive to certain colors. After doing a survey on 3000 jewelry images randomly on different online shopping websites (ebay.com, amazon.com, flipkart.com), the error rate was found to be 14.8%. That is, the color or the type specified in the title of the product did not match the product. These kinds of error if occur regularly, it may affect the sellers as it will result in low customer satisfaction and hence, low buyers. The aim is to introduce intelligent systems in this area to make it more efficient and accurate.

1.2 Literature Review

There are numerous practices to achieve image classification [8]. Popular approaches include the use of Artificial Neural Networks [9, 10, 11], Nearest neighbor, Support vector machines [12, 13], Fuzzy measure [14] etc. There are many other works that show the use of image classification techniques in jewelry industry. For example, they were used to detect and classify damages on the jewelry stone [15].

To get categories classified we need feature extraction. Every image has information associated with it and the work of a feature extractor is to obtain that relevant information. There are various techniques to extract features, such as, speeded up robust features (SURF) [16], where the image perception is achieved with the help of a SURF descriptors [17, 18] and then a classifier is applied. SURF or speeded up robust features is a licensed algorithm which is used for object detection tasks. Another one is Gabor filter [16] named after Dennis Gabor. The filter is used for texture inspection and generates texture features from the inputs. The next filter is known as Principal components analysis (PCA) [16, 19] the role of which is to reduce the number of random variables presented in a dataset and hence, it exposes strong patterns that are presented in an input.

2. Related Theory

The dataset in this work consists of five various categories of jewelry (earrings, bracelets, necklaces, pendants and rings) which are further divided into 5 different batches of 1000, 2500, 5000, 7500 and 10,000 image, respectively. The main purpose of this work is to introduce techniques that are essential for the image classification. The first two methods use transfer learning mechanism for the training and evaluation of the dataset and the third experimentation is done from scratch. The first approach is carried out using a multiclass linear support vector machine (SVM) along with the features that were extracted from our input images using AlexNet. The second one follows the Inception-V3 model. Our techniques for classifying jewelry into their distinct categories are implemented using a traditional path where classifiers are trained by using the features which were extracted from the database images.

For example, when classification is done with the bag of features [20, 21] it makes the use of the features that were obtained using the SURF feature extractor to train along with a multiclass SVM [22]. The dissimilarity is that this work uses convolutional neural network to obtain distinct characteristics from the input data.

2.1 Convolutional Neural Networks

Kunihiko Fukushima in 1980 invented “neocognitron” [23] which is a deep CNN. He paved the way for further researches in this area. Other neural inspired models are HMAX [24] and LeCun98 [25], to name a few. The CNNs are continuously used in many different areas, varying from medical to industrial and are also very useful in identifying faces, objects etc. like in robots and self-driving cars. The layers of CNN [26, 27] can be of three types:

- Convolutional: This layer consists of a rectangular grid of neurons and requires that the previous layers be of the similar grid. The inputs in this section are taken from the previous layer. The weights of the neuron in this section are equivalent to the previous ones. So, this layer is just the image convolution of the previous layer.
- Max-pooling: After each convolutional layer, there is a pooling layer. The pooling layer produces a single output by taking the rectangular blocks from the convolutional layer and then subsamples it. The pooling operation can be of three types: Max pooling, Avg pooling and Sum pooling. These are just the linear combination of the neurons present in a particular window. The pooling layer for this work will always be max-pooling.
- Fully-connected: The output received from previous layers consists of preeminent features which are then fed into the fully-connected layer for further analysis. The term fully-connected indicates that all the neurons in the network are interconnected with each other. Below are some equations and a figure (Fig. 1) regarding the Convolutional neural networks [28].

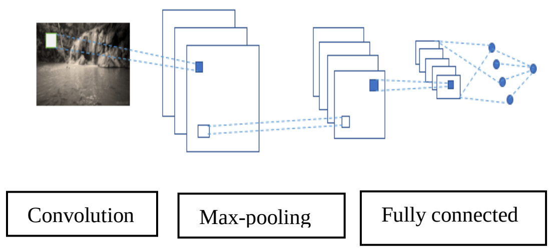

Fig. 1. Convolutional Neural Network

Cross-correlation: Suppose, there is an input image Y and a filter F of dimension $k \times k$, a cross correlation operation leading to an output image C can be defined as Eq. (1) and (2):

$$C = Y \otimes F \quad (1)$$

$$C(m, n) = \sum_{a=0}^{k-1} \sum_{b=0}^{k-1} \times (m + a, n + b) F(a, b) \quad (2)$$

Convolution: Input image Y and a filter F of dimension k × k, a convolution operation leading to an output image C can be defined as Eq. (3) and (4):

$$C = Y * F \quad (3)$$

$$C(m, n) = \sum_{a=0}^{k-1} \sum_{b=0}^{k-1} \times (m - a, n - b) F(a, b) \quad (4)$$

Over the past years, this technique has gained popularity due to their remarkable results in various fields ranging from self-driving cars, photo tagging applications to medical diagnosis. Therefore, a lot of researchers has explored this network to outperform the classical approaches for object recognition and detection. In this work, for the first approach, we have combined Convolutional neural network with the support vector machine. This can be achieved in two ways:

- The first way is to simply remove the fully-connected layer from a typical convolutional neural network and add linear support vector machine instead.
- The second method is to train the convolutional neural net in the usual way using the fully-connected layers. And when it is trained, the fully-connected layers are eliminated, and the features extracted by the convolutional layer are supplied back into the SVM for further training.

For the second approach Inception-v3 model has been used and the model is trained on the same set of data. Inception-v3 is a CNN model which was recently released by Google along with TensorFlow.

2.2 Support Vector Machine

Support vector machine (SVM) [29, 30, 31] is a classifier that is used for comparisons and classification. It focuses on the points that are very hard to differentiate. SVM can be linear or non-linear depending upon the separation problem and separates inputs into various categories by using a separating line. SVM stands out from the other classifiers as it focuses only on the points that are most difficult to tell apart. Fig. 2 shows the diagram for support vector machine.

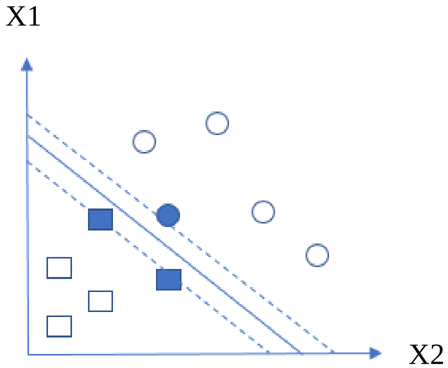

Fig. 2. Support Vector machine

If a line is passing close to the points then it is going to be inadequate because it will be sensitive towards noise that is present in the dataset and therefore, it will not be able to conclude properly. So, the target is to locate the line that is crossing as far as possible from all the points. This paper used multiclass linear support vector machine [32] to achieve the classification.

2.3 Inception-V3

Inception-v3 is a recently released model along with TensorFlow [33] by Google. Inception-v3 [34] model uses inception modules, which consists of different sub-models inside a bigger model. The architecture is that of convolutional neural networks but instead of having one operation per layer, it consists of combination of operations for every layer. At each layer of our ConvNet we must decide as to what type of convolution we want to make. Fig. 3 explains the basic architecture of the inception model.

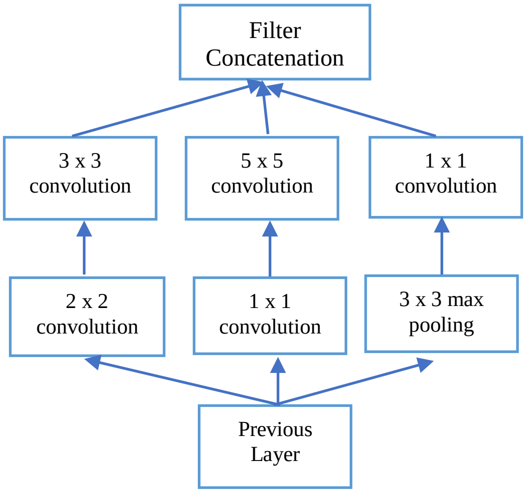

Fig. 3. Basic Inception model architecture

2.4 Transfer Learning

Transfer learning is the process of transferring knowledge from one learned task to the new task in machine learning. This is used because commonly ConvNet is trained on a very large dataset, therefore, by using transfer learning the knowledge of a learned feature of a pre-trained network is transferred to the new problem. Training a convolutional network from the scratch is not easy as it is a very intensive process that requires days or even weeks. That is why transfer learning is done to transport the knowledge from a network which is usually trained on a large dataset to a new network. In this process, the initial layers of the pre-trained network are generally fixed, only the last few layers are adjusted according to the need of the new dataset to learn other specific features. The approaches used in this work also contains transfer learning methodology.

3. Material and Methods

This section covers the techniques that are used in our work, which is, Convolutional neural networks along with the support vector machine and the Inception-v3 model.

3.1 CNN with Support Vector Machine

Convolutional neural network is composed of neurons like the ones in our brain which has the capability to learn and flourish. They are similar to the usual neural network, the difference is in their architecture. The CNN architecture is complex because of numerous layers and they take images as their inputs [26]. It makes the forward function much more effective to implement and helps in reducing the amount of variables that are being used in the network. Fig. 4 and Fig. 5 show different neural networks.

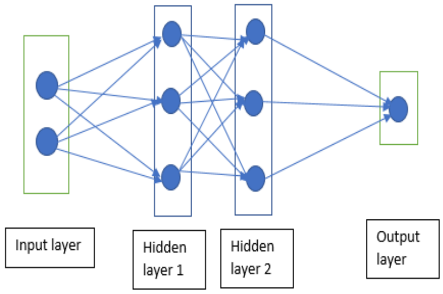

Fig. 4. A regular neural network

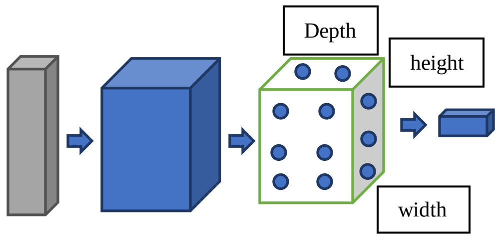

Fig. 5. Convolutional Neural network

Here the task is the image category classification of jewelry. For this approach, the last layer of the CNN was replaced with an SVM. The scheme is to use the

Convolutional neural network to obtain various features from the input images, then, by using the multiclass linear SVM train those features to classify them into their specific categories. To be more precise, the feature extraction was done with a pre-trained CNN, known as AlexNet. AlexNet was selected to reduce the time consumption as it contains plenty of filters to extract distinct features from the dataset images.

Support vector machine (SVM) was originally a two-class classifier but when it comes to the problems that are related to the real world often more than two categories are required. Due to which a multiclass linear SVM is used in this study.

3.1.1 Loading and Pre-processing images

The jewelry was distributed into five main categories (earrings, bracelets, necklaces, pendants and rings). Furthermore, each category was divided into five different batches of 1000, 2500, 5000, 7500 and 10,000 images respectively. The image category classifier was trained to distinguish between the 5 given categories. Equal number of images were considered so to maintain the balance of the training set. AlexNet, which is a pre-trained CNN feature extractor was used here. Fig. 6 shows the jewelry categories that were presented in our work.

Fig. 6. Five categories for training

3.1.2 Training and testing the images

The data was divided into training set and validation set. We selected 30% of images for the training data and 70% of images for the validation data randomly. By using AlexNet, all the features from the input images were extracted and trained alongside multiclass linear SVM classifier. The test features were then passed to measure the accuracy of the trained classifier and it was then applied to categorize new input images.

3.2 Inception V3 model

The next approach was the Inception-v3 [34]. It is a model for the image recognition system which consists of two parts. First one is the feature extraction with convolutional neural networks and the second part consists of classification using fully-connected and softmax layers. The pre-trained inception-v3 model achieves the accuracy for the recognition of general objects with thousands of classes. For this work, we utilized the last layers of the given model and trained the network according to the dataset.

3.3 Training from scratch

After the successful implementation of Inception V3 model, it was again selected to train the dataset from scratch. For this approach, a lightweight high-level API of TensorFlow [35, 36] called TF-slim was considered which is generally used for training and evaluating complex models. TF-slim is available as tf.contrib.slim. For each dataset, the image data was converted into TensorFlow's native TFRecord format. The format conversion is required when using TF-slim for image classification. The TFRecords are made as the conclusive version for the training purpose. The dataset was again divided into two parts: Training set and Validation set. 70% of the data was used for training and 30% was used for validation purpose.

The hardware used in this work consisted of a 64 bit operating system along with an 8 GB RAM, i7 processor and Nvidia GeForce GTX 1080 card. We used python 3.4 with TensorFlow 3.5 for our methods.

4. Result and Discussion

The data that was used in this work consisted of five different categories of jewelry (earring, necklace, ring, bracelet and pendant). The classification was tough because all the images had distinct backgrounds as they were taken directly from the jewelry industries and other big companies. There were plenty of images and all of them were in JPEG format. The objective in the beginning was to use the approaches that requires transfer learning technique to categorize jewelry images and then to use one of those technique to train the dataset from scratch.

4.1 CNN with SVM

It was observed that with the increase in the number of images the classification accuracy also increased as shown in the Table 1.

Table 1. Classification accuracy of 5 categories (SVM).

	Earrings	Necklace	Rings	Bracelets	Pendants
	For 1000 images = 99.19% in 3 minutes				
Success%	98.70%	96.60%	98.80%	96%	97.30%
Time	52 sec	52 sec	51 sec	52 sec	53 sec
	For 2500 images = 97.04% in 6 minutes				
Success%	99.28%	95.52%	99.24%	96.76%	97.20%
Time	9 min	2 min	2 min	3 min	2 min
	For 5000 images = 98.29% in 10 min				
Success%	99.06%	98.56%	99.46%	96.60%	99.26%
Time	4 min	5 min	5 min	4 min	6 min
	For 7500 images = 98.70% in 15 minutes				
Success%	99.28%	97.92%	99.33%	98.65%	99.17%
Time	7 min	6 min	6 min	6 min	7 min
	For 10K images = 98.31% in 17 minutes				
Success%	99.30%	95.84%	99.31%	98.84%	99.35%
Time	9 min	8 min	9 min	8 min	8 min

The result acquired was represented as a graph for an easy and accurate perception. Fig. 7 shows accuracy for the 5 categories.

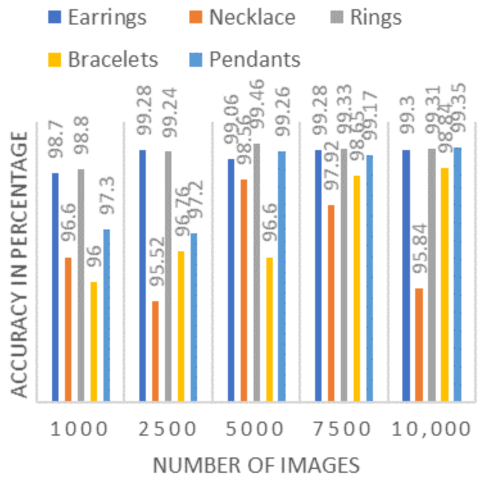

Fig. 7. Five category accuracy graph

4.2 Inception V3 model

In inception model too, with the increase in the number of images, the accuracy also increased. The result can be seen through Table 2.

Table 2. Classification accuracy of 5 categories (inception v3).

Earrings	Necklace	Rings	Pendants	Bracelets
	For 1000 images = 99.04% in 16 min			
99.40%	99.00%	99.60%	98.20%	99.00%
	For 2500 images= 99.10% in 17 min			
99.76%	99.20%	99.68%	98.96%	97.92%
	For 5000 images=99.28% in 20 min			
99.72%	99.12%	99.88%	99.28%	98.44%
	For 7500 images= 99.49% in 19 min			
99.62%	99.38%	98.45%	100%	100%
	For 10,000 images = 99.06% in 20 min			
99.56%	99.09%	99.55%	98.88%	98.26%

Fig. 8 shows the accuracy graph for the Inception-v3 model.

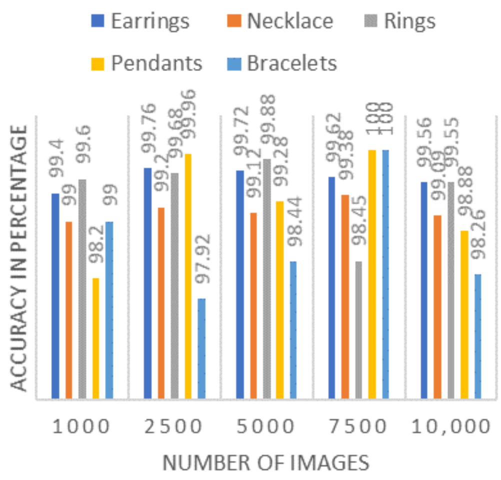

Fig. 8. Five category graph

For all the above batches, the errors were documented and represented as a confusion matrix. The errors took place because some of the images were misclassified into the other categories. The table below exhibits the confusion matrix for the batch of 10,000 images. The total number of images in table 3 is 3000 because the testing set had 30% data.

Table 3. Confusion matrix for 10K batch

	Bracelets	Earrings	Necklaces	Pendants	Rings	Total
Bracelets	2945	2	50	3	0	3000
Earrings	10	2966	16	4	4	3000
Necklaces	18	3	2975	2	2	3000
Pendants	6	25	3	2964	2	3000
Rings	6	1	3	4	2986	3000

The confusion matrix was prepared for the batch of 10,000 images. The error in the process took place because of the similarity between images.

Fig.9. Bracelet images miscalculated as necklaces.

Fig. 9 consists the images of bracelet that were misjudged as necklaces because of their resemblance. These kinds of similarity can even fool a human eye. They appear like a necklace and therefore, it is hard for any system to spot the difference. For this reason, the errors occurred in our work. In the future, we will try to find a solution for this problem.

4.3 Results obtained from scratch

Overall performance got better when the model was trained and evaluated from scratch. Dataset of raw images were changed into TFRecord format which was then trained and evaluated using Inception-v3 model on TensorFlow. The combined accuracy for this method was 99.13%. Below is Table 4 showing the number of images along with their accuracies And Fig. 10 shows the graph for the same table.

Table 4. Accuracy obtained when trained from scratch

Bracelets	Earrings	Necklace	Pendants	Rings
	For 1000 images = 99.33%			
99.00%	99.00%	99.66%	98.35%	99.66%
	For 2500 images = 99.60%			
99.60%	99.60%	99.60%	99.60%	100%
	For 5000 images = 98.80%			
99.53%	99.73%	99.68%	99.42%	99.73%
	For 7500 images = 99.06%			
99.51%	99.76%	99.68%	99.42%	100%
	For 10,000 images = 98.84%			
98.16%	98.86%	99.06%	98.70%	99.56%

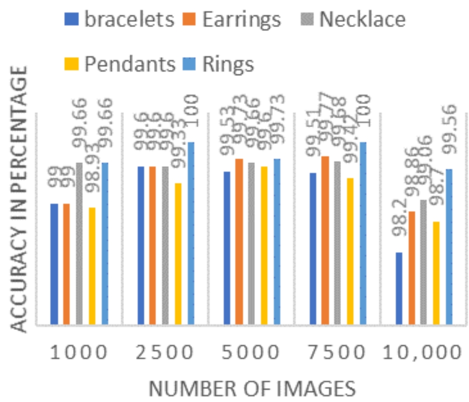

Fig.10. Accuracy when trained from scratch

4.4 Performance

The aim was to show different image classification approaches for the jewelry categorization. The methods opted in this paper were new for the jewelry industry. We made the use of transfer learning technique with SVM and Inception-v3 method to classify various categories of jewelry. Their accuracies and errors were recorded for the comparison. Both methods were good. However, Inception-v3 surpassed the SVM method by 0.9%. The comparison can be seen by their average accuracy graph shown below in Fig. 11.

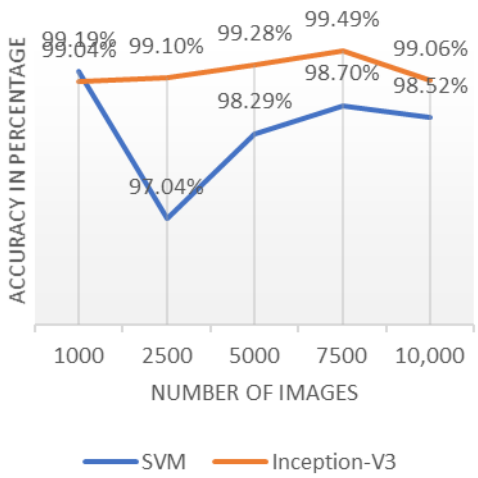

Fig.11. Average accuracy graph

Because Inception-v3 performed better performance, the latter was selected to use it to train the images from scratch. Firstly, the raw images were converted into TensorFlow's native TFRecord format for each of the 5 categories. This technique provided us with the overall accuracy of 99.13% and a very good consistency. The variation between the accuracies was between 98 to 100 (%), whereas in the transfer learning techniques it was between 95 to 100 (%).

The result can be seen by the graph shown in the Fig. 12.

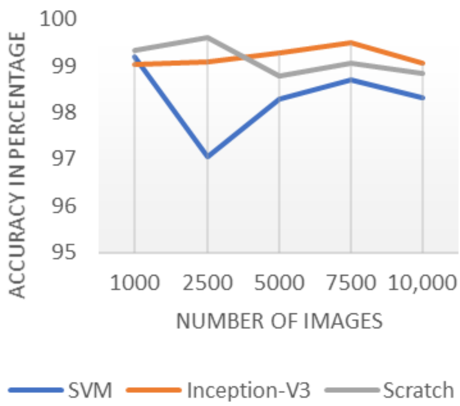

Fig. 12. Accuracy of all the three methods

5. Conclusion

In this work, we have suggested different approaches to categorize jewelry items but instead of using the traditional artificial neural nets we have used the convolutional neural networks as a substitute. The accuracies recorded exhibits how good the methods were. In the beginning, we selected two different classification techniques (SVM with CNN and Inception-v3) that were said to give the best results. Both were used on the same set of data to determine which technique will yield better results. Therefore, by evaluating the accuracies and the time taken by each of the technique it was found that Inception-v3 surpassed the other approaches by slight difference.

Inception-v3 yielded 99.19% accuracy, while SVM gave 98.30%. For the further work, Inception-v3 was selected alongside TensorFlow, where the model was trained from the scratch. The dataset was converted into the TFRecord format. Once, the conversion for each of the batch was done, we trained and tested our third approach. This method evaluated with the overall accuracy of 99.13%.

For the future work, the objective is to improve the model by adding more attributes (color, type, shape) which will help the model in giving the detail description of a particular image.

Acknowledgement

We would like to thank the Faculty of Engineering (Thammasat University), for their valuable support and resources.

References

[1] Fausett L. Fundamentals of Neural Networks architectures, algorithms and applications. Pearson Education Inc.; 1994.

[2] Jackson T and Beale R. Neural Computing: An Introduction. Bristol and Philadelphia: Institute of Physics Publishing; 1990.

[3] Chester M. Neural Networks: A tutorial. New Jersey: PTR Prentice hall; 1993.

[4] Pala Z, Yamli V and Ünlük IH. Deep Learning researches in Turkey: An academic approach. XIIIth International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH) 2017; 120-123.

[5] Medium.com [Internet]. MIT 6.S191: Introduction to Deep Learning. [cited 2017 Apr 6]. Available from: https://medium.com/@mit_intro_to_deep_learning

[6] Dauriz L, Remy N and Tochtermann T. A multifaceted future: The jewelry industry in 2020 [Internet]. [cited 2014 Feb] Available from: https://www.mckinsey.com/industries/retail/our-insights/a-multifaceted-future-the-jewelry-industry-in-2020.

[7] Amazon services [Internet], Available from:https://services.amazon.com/content/sell-on-amazon.htm/ref=sc_us_soa_strip?ld=SCSOAStriplogin

[8] Kamavisdar P, Saluja S and Agrawal S. A survey on image classification approaches and techniques. International Journal of Advance Research in Computer and Communication Engineering 2013; 2(1):1005-1009.

[9] Vasundhara DN and Seetha M. Rough-set and artificial neural networks-based image classification. 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I) 2016; 35-39.

[10] Ahmed SA, Dey S and Sarma KK, Image texture classification using Artificial Neural Network (ANN). 2011 2nd National Conference on Emerging Trends and Applications in Computer Science 2011; 1-4.

[11] Mahmon NA and Ya'acob N. A review on classification of satellite image using Artificial Neural Network (ANN). 2014 IEEE 5th Control and System Graduate Research Colloquium 2014;153-157.

[12] Thai LH, Hai TS and Thuy NT. Image classification using support vector machine and artificial neural network. Information Technology and Computer Science 2012;4(5): 32–38.

[13] Swiniarski RW and Hargis L. Rough sets as a front end of neural-networks texture classifiers. Neurocomputing 2001; 85-102.

[14] Chen DG, He Q, and Wang XZ. FRSVMs: Fuzzy rough set-based support vector machines. Fuzzy Sets System 2010; 161:596–607.

[15] Perfilieva I, Hodáková P, Vajgl M and Daňková M. Classification of damages on jewelry stones: Preprocessing. Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS) 2013; 783-788.

[16] Jain A, Srivastava S and Soman S. Transfer learning using adaptive SVM for image classification. IEEE Second International Conference on Image Information Processing (ICIIP-2013) 2013; 580-585.

[17] Bay H, Tuytelaars T, and Van Gool L. Surf: Speeded up robust features. In European Conference on Computer Vision 2006.

[18] Bay H, Tuytelaars T, Van Gool L and Ess A. Speeded-Up Robust features (SURF). Computer Vision and Image Understanding 2008; 110(3):346-359.

[19] Lai Z, Xu Y, Chen Q, Yang J and Zhang D. Multilinear Sparse Principal Component Analysis. IEEE Transactions on Neural Networks and Learning Systems 2014; 25(10):1942-1950.

[20] Grzeszick R, Plinge A and Fink GA. Bag-of-Features Methods for Acoustic Event Detection and Classification. IEEE/ACM Transactions on Audio, Speech, and Language Processing 2017; 25(6):1242-1252.

[21] Kamal I and Oubaha J. Car recognition using the bag of features method. 2016 5th International Conference on Multimedia Computing and Systems (ICMCS), Marrakech 2016; 99-102.

[22] Wang Z and Xue X. Multi-Class Support Vector Machine 2014; 23-48.

[23] Fukushima K, Miyake S and Ito T. Neocognitron: A neural network model for a mechanism of visual pattern recognition. IEEE Transactions on Systems, Man, and Cybernetics 1983; SMC-13(5):826-834.

[24] Serre T, Wolf L, Bileschi S, and Riesenhuber M. Robust object recognition with cortex-like mechanisms. IEEE Trans. Pattern Anal. Mach. Intell. 2007; 29(3):411-426.

[25] LeCunn Y, Bottou L, Bengio Y and Haffner P. Gradient-Based learning applied to document recognition. Proceedings of IEEE 1998; 86(11): 2278-2324.

[26] Stanford.edu [Internet]. Available from: http://ufldl.stanford.edu/tutorial/supervised/ConvolutionalNeuralNetwork

[27] Medium.com [internet]. Convolutional Neural Networks (CNNs or ConvNets). [cited 2017, Sep 28]. Available from: https://medium.com/@phidaouss/convolutional-neural-networks-cnn-or-convnets-d7c688b0a207

[28] Karpathy A and Johnson J. CS231n: Convolutional Neural Networks for Visual Recognition [Internet]. [cited 2016 Jan] Available from: http://cs231n.github.io

[29] Wikipedia. Support Vector Machines [Internet]. [cited 2017 Dec 23]. Available from:https://en.wikipedia.org/wiki/Support_vector_machine

[30] Cortes C and Vapnik V. Mach Learn 1995; 20: 273-297.

[31] Ben-Hur A, Horn D, Siegelmann H and Vapnik V. Support Vector Clustering. Journal of Machine Learning Research 2001; 2: 125-137.

[32] Hsu CW and Lin CJ. A comparison of Methods for Multiclass Support Vector Machines. IEEE Transactions on Neural Networks 2002; 13(2):415-425.

[33] Wikipedia. TensorFlow [Internet]. [cited 2018 Jan 8]. Available from: https://en.wikipedia.org/wiki/TensorFlow

[34] Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z. Rethinking the Inception architecture for computer vision. arXiv preprint, 1512.00567 2015. arxiv.org/abs/1512.00567.

[35] Datacamp.com [Internet]. [cited 2017 July 13]. Available from: https://www.datacamp.com/community/tutorials/tensorflow-tutorial

[36] Tensorflow.org [Internet]. [cited updated 2018 Mar 29]. Available from: https://www.tensorflow.org/tutorials